WISKUNDIGE ASPECTEN VAN HET NEDERLANDSE ADMINISTRATIENUMMER VOOR PERSONEN

door dr. J. Verhoeff

Lezing gehouden voor het Nederlands Rekenmachine Genootschap op 12 december 1969 te Amsterdam

INLEIDING

Het Nederlandse administratienummer is een redundante decimale code, met zeer goede foutdetectie mogelijkheden. Deze zal in dit artikel kortheidshalve Dupoco genoemd worden.
De belangrijkste eigenschappen van deze code zijn:
De codewoorden, die alle juist uit 10 cijfers bestaan, zijn dusdanig gekozen dat er nooit twee gelijke cijfers naast elkaar in een nummer voorkomen.
Alle codewoorden (nummers) voldoen aan twee (vaste) modulo 11 checkvergelijkingen, wat o.a. tot gevolg heeft dat fouten in één der cijfers gecorrigeerd kunnen worden.
Tenslotte is het een zogenaamde niet-classificerende code, m.a.w. er wordt geen relatie tussen karakteristieke eigenschappen van een persoon en het hem toegevoegde nummer aangebracht. Dit is bereikt door de nummers in een (pseudo) willekeurige volgorde uit te reiken. In feite is deze volgorde niet zo zeer aselect, maar er is sprake van een volgorde, die ontstaat door de natuurlijke (lexicografische) volgorde te versluieren en wel dusdanig dat het praktisch onmogelijk is met de hand (of het hoofd) het proces om te keren. Het betekent dus niet dat één van de talrijke mogelijke volgordes willekeurig is gekozen, doch veeleer dat de gekozen volgorde „zo min mogelijk" gelijkenis vertoont met de lexicografische. Een betere term dan pseudo willekeurig zou dan ook zijn: onordelijk of chaotisch.
In deze lezing zullen slechts de wiskundige aspecten van de structuur en de constructie van de code besproken worden. De redenen waarom een code van juist deze structuur gewenst geacht werd zullen niet ter discussie gesteld worden.

1 De structuur van de dupoco

De eisen, gesteld door een voorbereidende commissie, waren:

- a De code moet zuiver decimaal zijn.
- b De codewoorden moeten een vaste lengte hebben, d.w.z. alle nummers dienen evenveel cijfers te hebben.
- c De code moet tenminste 20 miljoen nummers hebben.
- d Er mogen geen gelijke cijfers naast elkaar in enig nummer voorkomen.
- e Hoogstens twee der cijfers mogen een checkcijfer zijn.
- f De codewoorden moeten in een vaste maar schijnbaar willekeurige volgorde uitgereikt worden. Het moet echter mogelijk zijn om die volgorde te reconstrueren met behulp van een algorithme, dat echter zo ingewikkeld moet zijn dat uitvoering zonder machinale hulpmiddelen praktisch uitgesloten is.

Deze eisen werden vertaald in de volgende zes voorwaarden:

1. De codewoorden bestaan uit alle 10 decimalen. Zij hebben dus de vorm $a_0a_1a_2a_3a_4a_5a_6a_7a_8a_9$, waarbij elke a_j een der cijfers 0 1 2 3 4 5 6 7 8 9 voorstelt met $0 \leqq j \leqq 9$.
2. Het meest linkse cijfer mag niet 0 zijn, dus $a_0 \neq 0$.
3. Naast elkaar staande cijfers mogen niet gelijk zijn, dus $a_i \neq a_i - a_{i-1}$ voor $1 \leqq i \leqq 9$.
4. De cijfers van elk codewoord voldoen aan:
$$\sum_{i=0}^{9} a_i = 0 \pmod{11}.$$
5. De cijfers van elk codewoord voldoen bovendien aan:
$$\sum_{i=0}^{9} 2^i a_i = 0 \pmod{11}.$$
6. De chaotische volgorde wordt in twee stappen bereikt, die in sectie 3 pas gedefinieerd zullen worden.

De bovenstaande voorwaarden zullen in het vervolg aangehaald worden als voorwaarde j, met j = 1, 2, 3, 4, 5, 6.

2 De fabricage van het dupoco

Na het zien van al deze voorwaarden zou men zich, met een zekere ongerustheid, af kunnen vragen of er nog wel codewoorden bestaan die er aan voldoen. In het vervolg zal echter blijken dat er 28816215 nummers in de code zijn waarmee uiteraard deze vraag beantwoord is.
Het is natuurlijk eenvoudig om alle 10-cijferige nummers te produceren, d.w.z. het proces is eenvoudig, maar het kan wel lang duren. Het is simpelweg tellen van 0 tot 9999999999. De eerste 10^9 nummers beginnen met een 0, te weten 0000000000 - 0999999999. Er blijft dus 90 % over als aan de voorwaarden 1 en 2 is voldaan. Hiervan vervalt weer 10 % omdat het tweede cijfer gelijk is aan het eerste enzovoorts, zodat tenslotte „slechts" 9^{10} woorden over blijven die aan de eisen 1, 2 en 3 voldoen. Dit is ongeveer 35 %.
Van deze uitgedunde nummercollectie zal gemiddeld 1 op de 11 voldoen aan de voorwaarde 4, terwijl ook 1 op de 11 aan voorwaarde 5 zal voldoen, zodat verwacht mag worden dat er ongeveer $9^{10}/121$ woorden aan de eisen 1 tot 5 zullen voldoen. Dit is ongeveer 28816400 of te wel 0,29 % van alle 10-cijferige decimale nummers.
Er zijn tenminste twee redenen om de bovenstaande sla-dood-methode niet te gebruiken, want primo hij is buitengewoon inefficient, secundo: het blijft bijzonder moeilijk om deze nummers doorelkaar te gooien, daar dit versluierproces de voorwaarden 1 tot 5 onveranderd moet laten en tertio: zij is zeer onelegant (1).
Er bestaat gelukkig een betere benaderingswijze die gebaseerd is op de navolgende opmerkingen, welke in sectie 3 nader gepreciseerd en bewezen zullen worden.

1 Noem de nummers die aan de voorwaarden 2 en 3 voldoen tweelingloze nummers en noem de nummers opgebouwd met symbolen, die 9 waarden aan mogen nemen, nonaire nummers. Het zal in de volgende sectie blijken dat de tweelingloze decimale nummers in een 1-1-duidige correspondentie gebracht kunnen worden met nonaire nummers met evenveel symbolen.

2 Ofschoon de restklassen modulo 9 met de optelling en de vermenigvuldiging als operaties geen lichaam vormen, bestaat er wel degelijk een lichaam met 9 elementen waarin zonder moeite omkeerbare transformaties kunnen worden gedefinieerd en uitgevoerd. Het is met behulp van een dergelijke transformatie dat de nummers doorelkaar geklutst zullen worden.

3 Het is mogelijk om de vergelijkingen van de voorwaarden 4 en 5 naar a_8 en a_9 op te lossen zodra a_0 t/m a_7 gegeven zijn. Dit berust op het feit dat de rekenkunde modulo 11 een lichaam vormt. Opgemerkt dient te worden dat deze oplossing modulo 11 is, zodat het voor kan komen dat de oplossingen de waarde 10 modulo 11 aannemen. In een dergelijk geval bestaan er geen corresponderende decimale cijfers a_8 en a_9 aangezien die slechts de waarden 0 t/m 9 modulo 11 hebben.

4 Uit de vorige opmerkingen volgt verder dat in ongeveer $(9/11)^2$ van de gevallen een gegeven tweelingloos 8-cijferig decimaal nummer kan worden uitgebreid tot een dito 10-cijferig nummer dat aan de voorwaarden 4 en 5 voldoet. Immers van de 11 mogelijke waarden modulo 11 voor a_8 zal 1 dier waarden verboden zijn doordat zij gelijk is aan het naastliggende cijfer a_7, terwijl de waarde 10 al bij voorbaat uitgesloten is. Iets dergelijks geldt voor a_9, met het kleine verschil dat de door a_8 uitgesloten waarde wel eens gelijk kan zijn aan de toch al verboden waarde 10.

De toegepaste procedure ziet er tenslotte uit als volgt:

stap 1: Breng in een lexicografische volgorde alle 8-cijferige nonaire nummers voort. Deze worden de nonaire basisnummers genoemd en aangeduid met $c_0c_1c_2c_3c_4c_5c_6c_7$.

stap 2: Kluts deze nonaire basisnummers doorelkaar met behulp van een (ingewikkelde) omkeerbare transformatie in het lichaam met 9 elementen. Deze nummers heten de versluierde nonaire nummers, zij worden aangeduid met $b_0b_1b_2b_3b_4b_5b_6b_7$, (2).

stap 3: Zet de versluierde nonaire nummers om in tweelingloze decimale nummers aangeduid met $a_0a_1a_2a_3a_4a_5a_6a_7$.

stap 4: Los de vergelijkingen uit de voorwaarden 4 en 5 op naar a_8 en a_9 in het lichaam van de restklassen modulo 11, met $0 \leqq a_8, a_9 \leqq 10$.

stap 5: Indien
$a_8 \neq 10$ en $a_9 \neq 10$ en $a_8 \neq a_7$ en $a_9 \neq a_8$, dan is er een goed codewoord ontstaan, dat dus uitgegeven kan worden. Zo niet dan is een goed bedoelde poging mislukt en men zal naar stap 1 terug moeten voor de volgende poging.

Op deze manier zijn 28816215 verschillende goede nummers gevonden in een schijnbare chaotische volgorde. De eerste 50 nummers zijn:

1234567853	4254560142	3853458937
8101267521	2976715304	5072374259
4370154234	9706713146	3913483184
5131292858	8904528469	2036743946
1430186280	5201256129	8309037248
9013472765	3267845875	9161285896
1639068173	7429064723	7589090746
3082328567	2491216314	5485653035
8697803248	2804573202	5670132171
4023490461	1927830473	6319058237
6982340183	2151235635	
7175603195	1382373123	
6785608654	3738963907	
4195635146	7291294037	
1514587238	6240146172	
2342392568	5360173658	
4561218510	9219092912	
3509026512	6594574301	
1868916565	4963464323	
3658943746	5728946021	

Het is zeer illustratief dat er in deze lijst van 50 nummers bij het zetten 4 fouten gemaakt werden, t.w. 3 maal 1 cijfer fout en 1 verwisseling van naburige cijfers. Al deze fouten worden de code ontdekt. De 3 enkelvoudige fouten kunnen automatisch verbeterd worden, terwijl de verwisseling als onverbeterlijk gesignaleerd wordt. Met een speciaal, in dit artikel niet genoemd, programma zou deze fout ook verbeterd kunnen worden.

3 Finesses van de fabricage

In deze sectie zullen de opmerkingen van de vorige paragraaf nader toegelicht worden.

ad 1: Laat $a_1a_2a_3 \ldots a_n$ een tweelingloos (decimaal) nummer zijn en stel $a_0 = 0$, dan zullen de cijfers $b_i = a_i - a_{i-1} \pmod{10}$, voor $1 \leqq i \leqq n$, alle ongelijk nul zijn. Derhalve vormen zij een nonair nummer. Ook het omgekeerde is waar: indien $b_1b_2 \ldots b_n$ een gegeven nonair nummer is, waarin de b_i's de waarden van 1 t/m 9 aan mogen nemen, dan zullen de cijfers a_i gedefinieerd door

$$a_i = \sum_{j=1}^{i} b_j \pmod{10}$$ een n-cijferig decimaal nummer vormen. Dit nummer is ten duidelijkste tweelingloos daar $a_1 = b_1 \neq 0$ en $a_i \neq a_{i-1}$ daar $a_i - a_{i-1} = b_i \neq 0$.

De boven beschreven procedure is uiteraard slechts één van de vele mogelijke om een dergelijke één-één-duidige correspondentie vast te leggen. Het voordeel van de gekozen procedure is dat zij zich met behulp van conventionele rekenmethodes (en dus ook rekenapparatuur) laat uitvoeren. Het rekenen modulo 10, afgekort (mod 10), wordt in het vervolg nog nader beschreven. Het is technisch bezien het rekenen in een accumulator met slechts één decimale positie.

ad 2: De begrippen groep en lichaam zijn alom bekend in de wiskunde, maar ten gerieve van de lezers met andere specialisatie zullen ze hier kort en onvolledig uitgelegd worden (3).

Losjes gezegd is een groep een verzameling G van elementen waarmee men kan optellen en aftrekken (zonder uit G te geraken). Meer precies is het een niet lege verzameling G, gesloten onder een binaire operatie $+$ (of $\times$) optelling (resp. vermenigvuldiging) genoemd, met de volgende eigenschappen:

De operatie voegt ondubbelzinnig aan elk tweetal elementen a en b van G een derde element c van G toe, deze relatie wordt aangegeven met $a + b = c$.
De operatie is associatief, dat wil zeggen
$(a + b) + c = a + (b + c)$ voor alle a, b, c uit G.
Er bestaat een zogenaamd eenheidselement e in G waarvoor geldt dat $a + e = e + a$ voor alle a uit G.
Tenslotte behoort er bij ieder element a uit G een element a′ uit G met de eigenschap, dat $a + a' = a' + a = e$, waarin e het eenheidselement van G is.
Geldt bovendien dat $a + b = b + a$ voor alle a, b uit G, dan heet de groep abels of ook wel commutatief. Het is gebruikelijk om bij abelse groepen de + als operatiesymbool te gebruiken en bij niet abelse groepen het ×-teken, dat dan dikwijls inconsequenterwijze wordt weggelaten.
Het gevolg van de bovenstaande eigenschappen is dat in elke groep de omgekeerde operatie, de aftrekking betekenis heeft. Bij niet-abelse groepen zijn er eigenlijk twee aftrekkingen, te weten één van links en één van rechts.
Een rijkere structuur heeft de ring, dit is een systeem met twee binaire operaties, + en ×, zodanig dat het systeem is gesloten onder elk van deze operaties, dat wil zeggen dat het resultaat van een optelling of van een vermenigvuldiging van een willekeurig tweetal elementen van het systeem, weer tot het systeem behoort. De elementen, met de optelling als operatie, behoren een abelse groep te zijn, terwijl bovendien de optelling en de vermenigvuldiging gekoppeld zijn door de zogenaamde distributieve wetten:
$a \times (b + c) = (a \times b) + (a \times c)$ en
$(a + b) \times c = (a \times c) + (b \times c)$.
Voor de vermenigvuldiging is slechts de associatieve wet gepostuleerd, dus $a \times (b \times c) = (a \times b) \times c$. Het is niet nodig dat de deling altijd ondubbelzinnig mogelijk is in een ring, dit in tegenstelling tot een lichaam, een nog rijker systeem gedefinieerd als volgt:
Een lichaam is een ring waarin bovendien geldt dat de elementen ongelijk nul, d.i. de eenheid van de optelgroep, een groep vormen met de vermenigvuldiging als operatie. In een lichaam is derhalve optelling, aftrekking, vermenigvuldiging en deling (behalve door nul) altijd mogelijk een ondubbelzinnig. Als het aantal elementen eindig is, dan spreekt men van een eindig lichaam of ook wel van een galois lichaam.
Men kan bewijzen dat in een eindig lichaam de vermenigvuldiging noodzakelijkerwijs commutatief is, dus $a \times b = b \times a$. Evenzeens kan bewezen worden dat het aantal elementen van een galois lichaam een macht van een priemgetal moet zijn. Ook kan bewezen worden dat er voor elke macht van een priemgetal altijd een lichaam bestaat met juist dat aantal elementen.
Dit heeft als onaangename consequentie dat er geen lichaam met 10 elementen bestaat, een grote slag voor de decimale code fans.
De definitie is (met opzet) zo, dat veel van de rekenkunde, die met de rationale getallen bedreven kan worden, doorgaat in elk lichaam. De theorie van de lineaire vergelijkingen, om slechts een onderwerp te noemen van belang voor het vervolg, gaat in zijn geheel door. Dit geldt niet voor sommige ordeningseigenschappen, zo is het zinloos om in het algemeen te beweren, dat men steeds grotere getallen krijgt wanneer men herhaaldelijk hetzelfde getal ergens bij optelt.
Eenvoudige voorbeelden van lichamen vormen de restklassen modulo een priemgetal p. Een restklasse modulo m is een verzameling gehele getallen die dezelfde rest geven na deling door m. De klassen kunnen gerepresenteerd worden door de niet negatieve gehele getallen k kleiner dan m, dus $0 \leqq k < m$.
Optelling of vermenigvuldiging van twee van dergelijke representanten resulteert in een geheel getal, dat uiteraard in een van de m mogelijke klassen terecht komt. De representant van die klasse wordt de som of het product modulo m genoemd. Zo is $8 + 7 = 4 \pmod{11}$ en $5 \times 9 = 1 \pmod{11}$ ook is $2 + 1 = 0 \pmod 3$ en $2 \times 2 = 1 \pmod 3$. Men kan bewijzen dat de deling modulo m slechts dan ondubbelzinnig mogelijk is als m ondeelbaar is. Voor $m = 10$ is dit geenszins het geval, daar bijvoorbeeld $2x = 1 \pmod{10}$ geen oplossing heeft, terwijl $2x = 4 \pmod{10}$ er twee en $5x = 5 \pmod{10}$ er vijf heeft. De oorzaak van deze kwaal is dat $2 \times 5 = 0 \pmod{10}$, reden waarom 2 en 5 nuldelers genoemd worden. Nuldelers verknoeien ten duidelijkste de nette rekenkundige eigenschappen. De restklassen modulo 10 vormen een ring, doch geen lichaam.
Zoals reeds eerder opgemerkt werd, is een lichaam met 9 elementen gebruikt om de nummers doorelkaar te klutsen.
De restklassen modulo 9 vormen evenmin een lichaam als die modulo 10, zo zijn 3 en 6 nuldelers, daar $3 \times 6 = 18 = 0 \pmod 9$. Aangezien echter 9 een macht is van een priemgetal moet er een lichaam met 9 elementen bestaan. De elementen van dit lichaam kunnen complexe ternaire getallen genoemd worden. De motivering voor deze naam volgt hieronder.
Er zijn precies negen complexe getallen van de vorm $a + bi$, met $a,b = 0, 1, 2$. Optelling en vermenigvuldiging van deze getallen kan op de gewone manier (d.i. in het lichaam van de complexe getallen of in de ring van de complexe gehele getallen) worden uitgevoerd, maar het resultaat moet soms worden teruggebracht in de uitgangsverzameling. Dit kan door zowel het reële als het imaginaire deel modulo 3 te reduceren. Zodoende wordt de som van $1 + 2i$ en $2 + 2i$, dus $3 + 4i$, na reductie $0 + 1i$.
Men kan ook zeggen dat het lichaam van de restklassen modulo 3 wordt uitgebreid met een wortel van $x^2 + 1 = 0$. Nog anders gezegd: de ring van de veeltermen over het lichaam van de restklassen modulo 3, gereduceerd modulo de veelterm $x^2 + 1$ is een lichaam met negen elementen. De clou is natuurlijk dat $x^2 + 1$ irreducibel (d.i. ondeelbaar) is modulo 3, of wat op hetzelfde neerkomt, dat $x^2 + 1 \neq 0 \pmod 3$ voor $x = 0, 1, 2$. Dit zou niet zo zijn als men modulo 5 werkte, aangezien $x^2 + 1 = (x + 2)(x + 3) \pmod 5$, of te wel $\sqrt{-1} = 2 \pmod 5$. Men had ook hetzelfde gekregen als men de complexe getallen van de vorm $a + b\sqrt{2}$, met $a, b = 0, 1, 2$, had genomen. Deze getallen vormen, onder de gewone complexe optelling en vermenigvuldiging na reductie modulo 3, ook een lichaam met 9 elementen en wel in principe hetzelfde lichaam.
Door nu op de één of andere vaste, doch overigens willekeurig te kiezen manier de negen elementen te coderen met de cijfers 1, 2, 3, 4, 5, 6, 7, 8, 9 is het mogelijk om een optelling en een vermenigvuldiging voor deze cijfers te definiëren, zodanig dat zij een gesloten rekensysteem vormen. In het bijzonder kan men dus lineair vergelijkingen opstellen en oplossen. Een transformatie

$b_i = \sum_{j=1}^{n} t_{ij}c_j$, voor $1 \leqq i \leqq n$, is omkeerbaar zodra de determinant $|t_{ij}|$ ongelijk 0 is (dan laten zich de c_j's oplossen als de b_i's gegeven zijn). Men merke op dat 0 hier het nulelement van het lichaam voorstelt, d.i. $0 + 0i$ (of $0 + 0\sqrt{2}$) met i de imaginaire eenheid, terwijl 0 nu uiteraard de restklasse van de door drie deelbare gehele getallen voorstelt. Het genoemde nulelement van het lichaam kan overigens met elk van de cijfers 1 t/m 9 gecodeerd zijn. In het geval van de dupoco is de volgende codering toegepast $0 + 0i \rightarrow 1$; $1 + 0i \rightarrow 2$; $2 + 0i \rightarrow 3$; $0 + 1i \rightarrow 4$; $1 + 1i \rightarrow 5$; $2 + 1i \rightarrow 6$; $0 + 2i \rightarrow 7$; $1 + 2i \rightarrow 8$; $2 + 2i \rightarrow 9$, zodat de tabellen voor de optelling en de vermenigvuldiging worden

+	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9
2	2	3	1	5	6	4	8	9	7
3	3	1	2	6	4	5	9	7	8
4	4	5	6	7	8	9	1	2	3
5	5	6	4	8	9	7	2	3	1
6	6	4	5	9	7	8	3	1	2
7	7	8	9	1	2	3	4	5	6
8	8	9	7	2	3	1	5	6	4
9	9	7	8	3	1	2	6	4	5

×	1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7	8	9
3	1	3	2	7	9	8	4	6	5
4	1	4	7	3	6	9	2	5	8
5	1	5	9	6	7	2	8	3	4
6	1	6	8	9	2	4	5	7	3
7	1	7	4	2	8	5	3	9	6
8	1	8	6	5	3	7	9	4	2
9	1	9	5	8	4	3	6	2	7

Voor het doorelkaar schudden van de nonaire basisnummers is de volgende transformatie gebruikt:

$$\begin{aligned}
b_0 &= 6c_1 + 8c_6 + 4c_7 \\
b_1 &= 9c_0 + 6c_2 + 5c_5 + 3c_6 + 8c_7 \\
b_2 &= 7c_0 + 9c_1 + 6c_3 + 4c_5 + 9c_6 + 3c_7 \\
b_3 &= 8c_0 + 7c_1 + 9c_2 + 6c_4 + 6c_5 + 9c_7 \\
b_4 &= 3c_0 + 8c_1 + 7c_2 + 9c_3 + 4c_5 \\
b_5 &= 3c_1 + 8c_2 + 7c_3 + 9c_4 + 4c_6 \\
b_6 &= 3c_2 + 8c_3 + 7c_4 + 9c_5 + 4c_7 \\
b_7 &= 9c_0 + 3c_3 + 8c_4 + c_5 + 8c_6
\end{aligned}$$

Door substitutie van de nonaire cijfers $c_0 \ldots c_7$ in de bovenstaande lineaire veeltermen worden de cijfers $b_0 \ldots b_7$ van het versluierde nonaire nummer direct gevonden.

De transformatie is verkregen door negen keer de eenvoudige schuiftransformatie $b_0 = c_1$; $b_1 = c_2$; $b_2 = c_3$; $b_3 = c_4$; $b_4 = c_5$; $b_5 = c_6$; $b_6 = c_7$; $b_7 = 6c_0 + 8c_5 + 4c_6$ uit te voeren, waarbij optelling en vermenigvuldiging steeds volgens de bovenstaande tabellen, d.i. in het lichaam van de complex ternaire getallen, wordt uitgevoerd.

Uit de laatste vergelijking volgt $c_0 = 6^{-1} b_7 - 6^{-1}8c_5 - 6^{-1} 4c_6 = 5b_7 + 2c_5 + 8c_6 = 5b_7 + b_4 + 8b_5$, zoals men gemakkelijk narekent, gebruik makend van de tabellen en van de vergelijkingen $b_4 = c_5$ en $b_5 = c_6$. Daar bovendien $c_1 = b_0$; $c_2 = b_1$; $c_3 = b_2$; $c_4 = b_3$; $c_5 = b_4$; $c_6 = b_5$ en $c_7 = b_6$ geldt, blijkt dat de schuiftransformatie omkeerbaar is. Hetzelfde geldt dan voor de negen macht van deze transformatie, d.i. de klutstransformatie zelf.

Het nut van de schuiftransformatie is enerzijds dat de omkeerbaarheid op een prettige wijze aantoonbaar (resp. gegarandeerd) is, maar anderzijds laat een dergelijke transformatie, of een macht ervan, een gemakkelijke technische realisatie toe.

ad 3: Zoals reeds is opgemerkt vormen de restklassen modulo 11 onder de optelling en vermenigvuldiging modula 11, een lichaam, daar 11 priem is. Dat a_8 en a_9 uit de beide vergelijkingen $\Sigma a_i = 0$ en $\Sigma 2^i a_i = 0 \pmod{11}$ opgelost kunnen worden, volgt het gemakkelijkst door het te doen. Stel $x = a_8$ en $y = a_9$ en stel $A = \sum_{i=0}^{7} a_i$ en $B = \sum_{i=0}^{7} 2^i a_i$, dan geldt $A + x + y = 0$ en $B + 2^8x + 2^9y = 0$, beide mod 11, iets wat in feite overbodig is om op te merken als de optelling en vermenigvuldiging in het restklassenlichaam worden uitgevoerd. Daar dan $2^8 = 256 = 23 \times 11 + 3 = 3$ en dus $2^9 = 2 \times 3 = 6$ geldt volgt, na eliminatie van x, dat $3y = 3A - B$ of wel, na vermenigvuldiging met 4 (d.i. 3^{-1}), dat $12y = y = A - 4B$ en dus $x = 4B - 2A$ geldt. Dit betekent dat er altijd voor a_8 en a_9 restklassen modulo 11 zijn te vinden zodat de voorwaarden 4 en 5 vervuld zijn. Het betekent bepaald niet dat er ook altijd twee (decimale) cijfers zijn te vinden die zulks doen, aangezien er geen decimaal bestaat die gelijk is aan 10 modulo 11.

ad 4: Uit de beschouwingen ad 3 volgt dat er twee decimalen voor a_8 en a_9 zijn te vinden zodra zowel $x \neq 10$ en $y \neq 10 \pmod{11}$ geldt. In dat geval zal er dus een nummer $a_0a_1 \ldots a_8a_9$ zijn dat aan de voorwaarden 4 en 5 voldoet. Dit verlengde nummer behoeft echter niet tweelingloos te zijn, zelfs als van een tweelingloos 8-cijferig nummer wordt uitgegaan. Dit is slechts het geval als ook nog geldt $x \neq a_7$ en $x \neq y \pmod{11}$. Hieruit volgt dat gemiddeld slechts 9 van de 11 mogelijke waarden voor x en y aanleiding geven tot een tweelingloze uitbreiding van een gegeven tweelingloos 8-cijferig nummer, zodanig dat aan de voorwaarden 4 en 5 is voldaan. De verwachting voor het aantal goede 10-cijferige nummers is dus $9^8(9/11)^2$, of wel 28816400.

Het juiste aantal blijkt 28816215 te zijn. Door de voorwaarden 4 en 5 te generaliseren tot $\Sigma a_i = s \pmod{11}$ en $\Sigma 2^i a_i = t \pmod{11}$ worden 121 verschillende codes, die ongeveer hetzelfde aantal woorden moeten hebben, vastgelegd. Het gekozen geval met $t = s = 0$ is blijkbaar niet behept met de meeste codewoorden, daar de som 9^8 moet zijn. Met behulp van recurrente betrekkingen, welke hier niet gegeven zullen worden, kan men deze aantallen zonder veel moeite (met een computer) uitrekenen. Het blijkt dan dat het geval $s = 1$ en $t = 0$ de meeste woorden heeft, namelijk 28818214. Achteraf bezien zou deze keuze dus bijna 2000 Nederlanders

meer toelaten, iets wat in verband met de toch al beperkte ruimte misschien niet eens wenselijk is.

4 Enige deugden van de dupoco

Het doel van coderen in administratieve processen is meestal, zo niet altijd, verhoging van de efficiëntie. Natuurlijke namen vereisen veelal 20 alfabetische posities en daarenboven vaak nog extra informatie, zoals de geboortedatum. Een 8-cijferig decimaal nummer is daarom hoogst efficiënt en voor de Nederlandse bevolking ook voldoende. Het is echter juist deze efficiëntie die de code kwetsbaar maakt voor fouten. De dupoco met zijn 10 cijfers, is vanzelfsprekend minder efficiënt, aangezien slechts 1 van de 347 mogelijke nummers wordt toegelaten. Ontzaggelijk veel lager ligt echter het percentage dat de familienamen vormen van alle mogelijke combinaties van 20 letters. Bovendien verschillen de nummers van de dupoco onderling op minstens drie plaatsen, zoals verderop bewezen zal worden. De eigennamen daarentegen lijken soms gevaarlijk veel op elkaar, iets waar Wolkers en Wolbers van getuigen, om van naamlopers maar niet te spreken.

Dit is dan ook de reden waarom de kunstmatige codes, ofschoon meer efficiënt, toch een betere bescherming tegen fouten kunnen geven. In dit licht zullen de verschillende eigenschappen van de code worden beschouwd.

ad 1: De vaste lengte van de codewoorden geeft bescherming tegen de veel voorkomende fouten van het weglaten (of toevoegen) van één of meer cijfers. Het geeft de code ook een zeker karakter, iets wat kan helpen om verwarring met andere codes, met een andere lengte, te voorkomen. De beperking tot een zuiver decimale code heeft als voordeel dat invoerapparatuur voor numerieke gegevens gegebruikt kan worden. Bovendien is het zo dat de gemengde of zuiver alfabetische codes aanleiding geven tot meer fouten, juist indien de mens een rol speelt bij het gebruik van de code (3). Met een zuiver alfabetische code zou men aan 7 symbolen voldoende hebben om het Nederlandse volk te coderen met dezelfde redundantie als de dupoco. Het aantal fouten zou echter veel groter zijn, zodat een hogere redundantie nodig zou zijn om deze fouten te detecteren, waardoor de code weer lager zou worden (zeg 8 à 9 symbolen). De winst van 1 of mogelijk 2 symbolen per woord zou dus gaan ten koste van meer fouten, die weliswaar automatisch ontdekt zouden worden maar toch ongewenst zijn.

ad 2: Deze eis ($a_0 \neq 0$) is een logisch gevolg van de eerste eis, daar veel apparatuur, die met een vaste woordlengte werkt, automatisch de kortere woorden van links af met nullen aanvult, zodat het verschil tussen bijvoorbeeld 00021 en 21 verloren gaat. Bovendien hebben veel mensen de neiging om nullen aan het begin van een nummer weg te laten, mogelijk omdat men op school geen rekenkunde met een vaste woordlengte geleerd heeft.

ad 3: Deze voorwaarde ($a_i \neq a_{i-1}$) is een poging om de brokkenmakers onder de codewoorden uit te sluiten. De ervaring heeft geleerd, dat twee of meer gelijke cijfers naast elkaar vaak aanleiding tot fouten geven. Weliswaar is het weglaten van één of meer van deze cijfers de meest voorkomende fout, die wegens eis 1 toch ontdekt zou worden, maar voorkomen is beter dan genezen. Dit is misschien een goede gelegenheid om op te merken dat het soort fouten waartegen de code in de eerste plaats moet beschermen de menselijke fouten zijn die gedurende de invoerfase van het proces op kunnen treden en niet de machinefouten die gedurende de overdracht en verwerking kunnen ontstaan.

ad 4: Deze beide eisen samen garanderen dat elk tweetal
en nummers op tenminste drie plaatsen verschillen.
ad 5 Dit volgt uit het feit dat twee nummers, die op 8 van de 10 plaatsen gelijk zijn, dat noodzakelijkerwijze ook op de beide overige plaatsen moeten zijn. Laat $a_0a_1 \ldots a_9$ en $a'_0a'_1 \ldots a'_9$ de beide nummers voorstellen en laten j en k met $0 \leqq j < k \leqq 9$, de indices zijn van de cijfers waarvan niet gegeven is dat zij gelijk zijn.

Dus $a_i = a'_i$ voor $i \neq j$ en $i \neq k$. Door Σa_i en $\Sigma a'_i$, die beide nul modulo 11 zijn, van elkaar af te trekken, krijgt men $a_j - a'_j + a_k - a'_k = 0$ (mod 11). Door hetzelfde uit te halen met $\Sigma 2^i a_i$ en $\Sigma 2^i a'_i$ ontstaat $2^j(a_j - a'_j) + 2^k(a_k - a'_k) = 0$ (mod 11). Uit deze vergelijkingen volgt direct dat $(2^k - 2^j)(a_k - a'_k) = 0$ (mod 11) en daar $j \neq k$ en $j \neq k - 10$ geldt dat $2^k - 2^j \neq 0$ (mod 11), zodat dus $a_k - a'_k = 0$ (mod 11) moet gelden. Hieruit volgt $a_k = a'_k$ en in verband met de eerste gelijkheid, $a_j = a'_j$, zodat de nummers identiek zijn.

De spil van dit bewijs is dat de gewichten uit de voorwaarde 5, te weten 2^i, voor $0 \leqq i \leqq 9$, alle verschillend zijn modulo 11, zoals in de onderstaande tabel wordt tentoongesteld.

i	0	1	2	3	4	5	6	7	8	9
2^i	1	2	4	8	16	32	64	128	256	512
2^i (mod 11)	1	2	4	8	5	10	9	7	3	6

Omkering van deze tabel geeft na herrangschikking van de kolommen:

k	1	2	3	4	5	6	7	8	9	10	2^i (mod 11)
lo (k)	0	1	8	2	4	9	7	3	6	5	i

Dit is de tabel van de logarithme met het grondtal 2, van de restklassen modulo 11. Deze logarithme zal om verwarring te voorkomen met lo worden aangeduid. Men merke op dat in de vergelijking $\mathrm{lo}(a \times b) = \mathrm{lo}(a) + \mathrm{lo}(b)$ het product modulo 11 genomen wordt, maar de som van de logarithme is modulo 10 te interpreteren. De functie lo geeft dan ook een isomorfie tussen de vermenigvuldiggroep van de 10 elementen ongelijk 0 van het lichaam van de restklassen modulo 11 en de (cyclische) groep van de optelling modulo 10.

Omdat elk tweetal nummers op minstens drie plaatsen verschillen, wordt de code een minimum afstand drie code genoemd. Deze eigenschap geeft de mogelijkheid om fouten in één cijfer te corrigeren. In de laatste paragraaf wordt verteld hoe dat in zijn werk gaat.

ad 6: De meeste codes worden ontworpen voor een groeiende populatie. De codewoorden worden aan de leden van zo'n populatie in een zekere

volgorde uitgereikt. Meestal is dit de lexicografische volgorde. Het voordeel hiervan is dat het reeds gebruikte deel van de code onder controle is, zodat fouten die tot nog niet uitgereikte nummers leiden automatisch ontdekt kunnen worden tijdens de invoerfase, in plaats van tijdens de verwerking door het zoeken van een niet bestaand gegeven (record). Het nadeel van die volgorde is dat nummers van een aaneengesloten serie in het algemeen aan de linkerkant gelijk zijn, waardoor fouten in de rechterhelft niet bedoelde nummers van dezelfde serie geven. Een (pseudo) aselecte volgorde van uitreiking geeft een betere verdeling van de serie of van het gehele gebruikte deel over de coderuimte en daardoor een betere foutbescherming.
Als een omkeerbare transformatie de aselecte volgorde bepaalt dan is het gebruikte deel van de code evenzeer onder controle als bij de lexicografische uitreiking.
Tenslotte schijnt er een politiek voordeel te schuilen in de chaotische volgorde, daar er dan geen eenvoudige methode is om uit het nummer conclusies te trekken over de plaats en het tijdstip van afgifte.

5 Foutdetectie met de dupoco

Een fout in een codewoord zal ontdekt worden als het verminkte woord geen betekenis geeft, hetzij op zichzelf, hetzij in de context. Een woord dat geen betekenis heeft is ongebruikt, zodat voor foutdetectie altijd redundantie nodig is. De grondgedachte van een foutdetecterende code is het dusdanig aanbrengen van de redundantie dat de kans dat door een fout in een gebruikt (zinvol) codewoord er een niet gebruikt (zinloos) codewoord ontstaat zo groot mogelijk is. Het geheim van een goede code is dat de (toegelaten) codewoorden als het ware gespreid zijn over de ruimte van alle mogelijke woorden. Afstand tussen twee nummers is het aantal fouten dat gemaakt moet worden om de nummers in elkaar over te voeren. Dit is ten duidelijkste afhankelijkheid van wat men een fout wenst te noemen. Zo kan een verwisseling van twee ongelijke cijfers opgevat worden als één of als twee fouten, iets wat van het grootste belang is voor het ontwerp van de code. Er zijn zelfs auteurs die de overgang van een 6 naar een 7 of een 5 als één fout opvatten, maar de overgang van een 6 naar een 8 als twee fouten. De bovenbedoelde minimum afstand 3 is t.o.v. de zogenaamde onbeperkte enkelvoudige fouten. Dit zijn fouten in één der cijfers waarbij elk cijfer in elk ander cijfer over kan gaan.
Het is goed dat men zich realiseert dat er bij het optreden van een fout drie gevallen kunnen worden onderscheiden. De drie gevallen corresponderen met de drie soorten woorden, te weten cijfercombinaties die niet tot de code behoren, d.w.z. die niet voldoen aan de voorwaarden 1 tot 5 in het geval van de dupoco. Ten tweede nummers die wel aan alle eisen voldoen (dus syntactisch juist zijn), maar die niet gebruikt zijn in de toepassing in questie, d.w.z. die niet uitgereikt zijn aan een bevolkingslid in het geval van de dupoco. Ten derde die nummers die aan de eisen voldoen en die ook zijn gebruikt voor het coderen van een bevolkingslid. De gevallen zijn derhalve:

1 Het foutieve woord behoort niet tot de code (is dus syntactisch onjuist) en is dus een onnet woord en valt als zodanig direct door de mand.

2 Het foutieve woord behoort weliswaar tot de code, is dus netjes, maar het is ongebruikelijk in de context doordat het (nog) niet uitgereikt is. Ook deze fouten worden, al dan niet direct, ontdekt. Zij zijn hinderlijk doch ongevaarlijk.

3 Het foutieve woord is een net woord, dat behoort bij een ander, niet bedoeld, lid van de populatie. Dit soort fouten is gevaarlijk, daar zij aanleiding kunnen geven tot verkeerde opdrachten e.d. Dergelijke fouten worden veelal, met schade en schande, buiten het systeem ontdekt, bijvoorbeeld doordat Oma een oproep krijgt voor de militaire dienst of doordat de verkeerde patiënt geopereerd wordt. De lezer verzinne zo mogelijk zelf andere voorbeelden.

Het laatste soort fouten is beslist onvermijdelijk zodra er meer dan één toegelaten codewoord nodig is, het is immers juist de keuzevrijheid die fouten mogelijk maakt. De opzet van de dupoco is zo dat fouten van de eerste en de tweede soort automatisch in de invoerfase ontdekt kunnen worden, hetzij met een algorithme in een computer, hetzij met een speciaal controle-apparaat gekoppeld aan de invoerapparatuur. De keuze tussen die mogelijkheden hangt af van het totale systeemontwerp. Er is nagenoeg niets bekend over de fouten die men bij de dupoco kan verwachten. De code is ontworpen in de veronderstelling, dat er geen zeer groot verschil zal zijn met de fouten die men tot nu toe in allerlei administraties heeft opgemerkt (4).
Dit foutpatroon is ruwweg als volgt: 20 % cijfers vergeten of toegevoegd; 60 % één cijfer fout; 15 % dubbele fouten, waarvan ongeveer 90 % verwisselingen zijn; 5 % meervoudige en random fouten. Bij random fouten lijkt het foutieve nummer in het geheel niet op het bedoelde nummer, zij het dat er vaak een semantisch verband tussen de nummers kan bestaan doordat bijvoorbeeld het goede nummer het gironummer en het foute nummer het telefoonnummer van dezelfde persoon is. Het is redelijk om te verwachten, dat het percentage random fouten bij de dupoco lager zal liggen dan bij een giro of bank. Het percentage echte drie- of meervoudige fouten is meestal zeer gering, zeg minder dan 1 %.
Er is dan ook nog een onbekend percentage fouten waarbij per ongeluk expres een niet bedoeld nummer correct wordt ingevoerd, zoals bij girale opdrachten voor kan komen als de nummers van twee begunstigden door een rekeninghouder zelf verwisseld worden. Dit soort fouten is uiteraard immuun tegen elke code, hoe redundant ook.
Het zal duidelijk zijn, dat er bij een minimum afstand drie code tenminste drie fouten nodig zijn in een codewoord om een fout van de derde soort te geven. Hieruit volgt dat alle enkelvoudige en alle dubbele fouten dus zeker in de eerste categorie vallen en dus ontdekt zullen worden. Tevens zal duidelijk zijn dat een enkelvoudige fout een woord geeft dat van het goede codewoord een afstand 1 heeft en dus van elk ander goed codewoord minstens een afstand 2. Hierdoor kan in principe het bedoelde codewoord teruggevonden worden. Alleen als er meer dan 1 fout is gemaakt zal het foute woord verder

van het bedoelde afdwalen en mogelijk een ander, niet bedoeld, codewoord benaderen, resp. raken. In het laatste geval ontstaat een fout van de derde categorie, terwijl in het geval dat een (niet bedoeld) codewoord tot op een afstand 1 is benaderd, een automatisch uitgevoerde correctie er triomfantelijk een fout bij zal maken.
Weliswaar kan het feit dat er gecorrigeerd is worden aangegeven, zodat er achteraf nog andere semantische controles uitgevoerd kunnen worden, doch het is een voordeel als de code er op berekend is deze gevallen te beperken. Dit is bij de dupoco in ruime mate het geval, zoals verderop uitvoerig uit de doeken gedaan zal worden.

Voor het ontdekken van de fouten van de tweede soort is het nodig om het gebied van de nonaire basisnummers corresponderend met de gebruikte nummers te kennen. Bij het afgeven van de nummers in eerste aanleg is het zelfs mogelijk om per (kleine) serie dit gebied te onthouden. Een dergelijk gebied is bepaald door twee nonaire nummers, r_1 en r_2 welke het grootste en kleinste nummer van dat gebied voorstellen.
Het detectieproces is recht-toe-rechtaan, men behoeft slechts een aantal testen, corresponderend met de voorwaarden 1 t/m 6, uit te voeren. Voor een computerprogramma is de volgorde van uitvoering in wezen niet ter zake doende, zij het dat men een zekere tijdwinst kan bereiken door de testen die gemiddeld de meeste fouten vangen het eerst te doen. In een speciaal controle-apparaat, zoals ontworpen voor de dupoco, verdwijnt dit probleem voor een deel, daar er dan allerlei testen parallel worden gedaan.

De testen zijn:

1 Is het aantal karakters gelijk aan 10 en zijn het alle decimalen?
2 Is het eerste cijfer ongelijk 0?
3 Zijn alle naast elkaar voorkomende cijfers ongelijk aan elkaar?
4 Is $\sum_{i=0}^{9} a_i = 0 \pmod{11}$?
5 Is $\sum_{i=0}^{9} 2^i a_i = 0 \pmod{11}$?
6 Ligt het nonaire basisnummer, behorend bij het aangeboden decimale nummer in het gegeven gebied?

Van al deze testen behoeft wellicht alleen de laatste toelichting. Het nonaire basisnummer kan als volgt teruggevonden worden: Laat $a_0 a_1 \ldots a_9$ het gegeven decimale nummer zijn. Door de verschillen $a_i - a_{i-1}$, met $0 \leqq i \leqq 7$ en $a_{-1} = 0$, modulo 10 te bepalen wordt het versluierde nonaire nummer $b_0 b_1 \ldots b_7$ gevonden. Dus $b_i = a_i - a_{i-1} \pmod{10}$ en $1 \leqq b_i \leqq 9$.

De laatste ongelijkheden gelden in de veronderstelling dat de testen 2 en 3 in orde waren, anders is test zes overbodig. Ontsluieren betekent dat de 8 klutstransformatie-vergelijkingen van paragraaf 3 naar c_0, $c_1, \ldots, c_7$ moeten worden opgelost. Door de transformatiematrix in het lichaam der complex ternaire getallen te inverteren wordt de volgende inverse transformatie gevonden:

$$
\begin{aligned}
c_0 &= b_1 + 6b_2 + 4b_3 + 9b_4 + b_5 + 7b_7 \\
c_1 &= 5b_0 + b_2 + 6b_3 + 4b_4 + 5b_5 + 3b_6 \\
c_2 &= 5b_1 + b_3 + 6b_4 + 4b_5 + 5b_6 + 3b_7 \\
c_3 &= 8b_0 + 5b_2 + b_4 + 8b_5 + 5b_7 \\
c_4 &= b_0 + 8b_1 + 5b_3 \\
c_5 &= b_1 + 8b_2 + 5b_4 \\
c_6 &= b_2 + 8b_3 + 5b_5 \\
c_7 &= b_3 + 8b_4 + 5b_6
\end{aligned}
$$

Men kan gemakkelijk door substitutie de juistheid hiervan controleren. Gemakkelijk betekent bepaald niet dat het weinig tijdrovend is, iets wat ook beslist niet door de commissie gewenst werd. De lezer oordele zelf in hoeverre de procedure ongeschikt is voor handbewerking.
Hetzelfde resultaat kan verkregen worden door negen keer de inverse schuiftransformatie $c_0 = 5b_7 + b_4 + 8b_5$; $c_1 = b_0$; $c_2 = b_1$; $c_3 = b_2$; $c_4 = b_3$; $c_5 = b_4$; $c_6 = b_5$; $c_7 = b_6$ uit te voeren. Het aldus gevonden nonaire basisnummer moet dan tussen r_1 en r_2 in liggen.
Als alle testen goed uitvallen dan is er geen fout ontdekt. Meestal zal dit zijn omdat het nummer goed is, soms echter zal het nummer toch fout (d.w.z. niet bedoeld) zijn. Dit zal op grond van de bovengenoemde verwachting van het foutpatroon, slechts in 65 van de miljoen fouten optreden. Hierbij wordt er van uitgegaan dat 13 miljoen nummers in gebruik zijn voor het Nederlandse volk, zodat de zesde test altijd nog 45 % (13000000 / 28816215) van de willekeurige fouten ontdekt. Neemt men aan dat ca. 1 % van de mutaties met een fout nummer is behept, dan zou dit neerkomen op 65 onontdekte fouten op de 100 miljoen mutaties. Bij de invoering van het nummer kan dit aanzienlijk verbeterd worden door de nummers in kleine (chaotische) series uit te geven, waardoor de zesde test veel effectiever wordt. Bij een serie van een miljoen zou weer een factor 13 gewonnen zijn. Een zuiver random code, gebruikt met dezelfde redundantie van 1 op 769 ($10^{10} / (13.10^6)$), zou 20 keer zoveel fouten laten ontsnappen. Facultatief kan, indien een fout is geconstateerd, een correctie ondernomen worden. Dit gebeurt in de veronderstelling dat er één cijfer fout is, het meest waarschijnlijke geval. Is deze veronderstelling juist dan wordt er goed gecorrigeerd, zo niet dan wordt er een fout bijgemaakt.
Zoals reeds opgemerkt werd, is het dus van belang om te ontdekken of dit het geval is. Als test 1 het nummer afwijst is het een duidelijke zaak.

Aan het resultaat van de testen 2 en 3 kan ook wel iets gezien worden, nl. als er teveel gelijke buurcijfers voorkomen is een enkelvoudige fout uitgesloten. Het is echter praktischer om achteraf het gecorrigeerde nummer weer aan deze testen te onderwerpen. Hiermee wordt hetzelfde resultaat bereikt, men zou één correctieslag kunnen besparen ten koste van een minder eenvoudige vorm voor de testen 2 en 3. Overigens moet het gecorrigeerde nummer natuurlijk toch nog opnieuw op tweelingen getest worden. Anders ligt het met de testen 4 en 5.
De beide controlevergelijkingen zijn zo gekozen, dat een enkelvoudige fout altijd beide vergelijkingen onklaar maakt. Op zichzelf zou dit 20/121 van de meervoudige fouten signaleren, maar doordat de vergelijking van voorwaarde 4 symmetrisch is in de cijfers, zullen alle verwisselingsfouten deze intakt laten. Alvorens met deze beschouwingen voort te gaan is het wenselijk om eerst de

correctie procedure uiteen te zetten. Laat $a_0a_1 \ldots a_9$ het te corrigeren nummer zijn en laat het cijfer a_j, met $0 \leqq j \leqq 9$ het foute cijfer zijn, waarvan de juiste waarde a'_j zij. Er zal dus gelden: $\sum_{i=0}^{9} a_i = A \not\equiv 0 \pmod{11}$ en $\sum_{i=0}^{9} 2^i a_i = B \not\equiv 0 \pmod{11}$, is immers A of B nul modulo 11 dan is het duidelijk dat er geen enkelvoudige fout gemaakt kan zijn en dan wordt er niets ondernomen wat toch mis is. Daar het correcte nummer aan de voorwaarden 4 en 5 voldoet, volgt dat $A \equiv a_j - a'_j \pmod{11}$ en $B \equiv 2^j (a_j - a'_j) \pmod{11}$ waaruit de relatie $B \equiv 2^j A \pmod{11}$ ontspruit. Dit is, door het nemen van de duale logarithme modulo 11, terug te brengen tot $j \equiv lo(B) - lo(A) \pmod{10}$, waardoor j bepaald is als één der waarden 0 t/m 9. De juiste waarde van a_j volgt dan uit $a'_j \equiv a_j - A \pmod{11}$. Dit kan alleen de uitkomst 10 geven als er geen enkelvoudige fout gemaakt was. Mocht er dus wel 10 uitkomen dan blijkt daardoor de onverbeterbaarheid van het nummer. Dit kan ook nog achteraf blijken wanneer het verbeterde nummer niet door de testen 2, 3 en 6 komt.
Een nummer wordt derhalve niet gecorrigeerd als:

1 Het aantal cijfers niet 10 is.
2 $\Sigma a_i \equiv 0 \pmod{11}$, dit gebeurt als er van een verwisselingsfout gemaakt is.
3 $\Sigma 2^i a_i \equiv 0 \pmod{11}$, dit gebeurt bij toeval (ca. 9 %) bij meervoudige fouten.
4 Als de correctieprocedure een waarde 10 modulo 11 aangeeft voor het verbeterde cijfer.
5 Als het gecorrigeerde nummer niet tweelingloos blijkt te zijn.
6 Als het nonaire basisnummer, behorend bij het gecorrigeerde decimale nummer niet in het gegeven nonaire gebied ligt.

De verbeteringsinrichting geeft naar verwacht wordt in ongeveer 0,1 % van de verbeterde gevallen een verkeerde uitkomst. In ongeveer 40 % van de gevallen zal de fout onverbeterbaar gesignaleerd worden, de helft van deze gevallen betreft fouten in de lengte van het nummer.
Voor ingewijden zij opgemerkt dat een perfecte foutcorrigerende code nooit de aardige eigenschap zou kunnen hebben van het niet (verkeerd) corrigeren van de verwisselingsfouten.
Het is mogelijk om een controle-apparaat te bouwen dat de fouten automatisch detecteert en desgewenst corrigeert zo dit mogelijk is. De details van de constructie worden bij deze gelegenheid niet besproken, hoewel er enige niet-triviale kunstgrepen bij betrokken zijn.

LITERATUUR:

1 Jean Dulieu
„Paulus en de drie rovers"
Amsterdam 1963

2 Berger Janssen
„Random number generators"
Stockholm 1966

3 Saunders Machane & Garrett Birkhoff
„Algebra"
New York 1953

4 J. Verhoeff
„Error detecting decimal codes"
Math. Center tract no.: 29
Amsterdam 1969

DE VASTLEGGING VAN GEGEVENS BIJ DE AUTOMATISCHE INFORMATIEVERWERKING VOOR DE MIDDELGROTE EN KLEINE ONDERNEMING*

door prof. A. J. van 't Klooster

Het is een verheugend verschijnsel, dat in het afgelopen jaar de automatisering in het algemeen en de mogelijkheden voor het middelgrote en kleine bedrijf in het bijzonder in de pers de nodige aandacht heeft gekregen. Enerzijds is deze gericht op het gebruik maken van computersystemen in eigen beheer, hetgeen mede door de ontwikkeling van kleine apparatuur ruimere mogelijkheden is gaan bieden en anderzijds op de ontwikkeling van service-mogelijkheden. Het is begrijpelijk, dat de moderne ontwikkeling van het gebruik van computers op afstand met behulp van eenvoudige in- en uitvoerapparatuur (terminals bij datatransmissie) veel aandacht heeft gekregen.
In Nederland zijn reeds toepassingen gerealiseerd, waarbij een groot aantal gebruikers met behulp van schrijfmachines, beeldschermen e.d., onmiddellijk toegang hebben tot een elders opgesteld groot computersysteem (time sharing). Bijzonder aantrekkelijk is daarbij, dat na de intoetsing van de variable gegevens op de schrijfmachine, de resultaten van de verwerking vrijwel onvertraagd op dezelfde schrijfmachine kunnen worden afgedrukt of op een beeldscherm worden geprojecteerd.
De toepassingsmogelijkheden richten zich echter nog

* Uit: Het Financieel Dagblad dd. 5 maart 1970.